

# Renyi-Ulam Games and Forbidden Substrings

NIKOLAI BELUHOV

**Abstract.** The Renyi-Ulam game is played between two players, the Seeker and the Obscurer. The Obscurer thinks of a number between 1 and  $n$ . The Seeker wishes to identify that number. On each turn, the Seeker asks the Obscurer whether her number belongs to a specific subset of the numbers from 1 to  $n$ . The Obscurer answers either yes or no, and her answer is either true or false. The series of truths and lies given by the Obscurer must conform to a restriction  $R$  that the players have agreed on in advance.

We give criteria on the restrictions  $R$  that allow the Seeker to win. Then we apply our results to the study of restrictions characterized by forbidden substrings. In particular, we give a complete classification of all such restrictions characterized by two forbidden substrings, elaborating on Czyzowicz, Lakshmanan and Pelc's classification of all such restrictions characterized by one forbidden substring.

## 1 Introduction

Renyi-Ulam games were first introduced by Alfred Renyi in 1961 in [3]. Stanislaw Ulam rediscovered the game and initiated its popularity in 1976 in [4].

In its original form, the Renyi-Ulam game is played between two players, a Seeker and an Obscurer, as follows. The Seeker seeks to identify an object, say, a number between 1 and  $n$ , known to the Obscurer. To this end, he may ask arbitrary yes-no questions. The Obscurer replies, for the most part, accurately – with the exception of at most one lie. What is the minimum number of questions that the Seeker needs to attain his goal, and what strategy should he follow?

Many variations on this basic theme have been explored in the literature. A thorough survey is [2].

We focus on Renyi-Ulam games in which the Obscurer's capacity to mislead the Seeker is restricted in a different way but the rules of the game are otherwise unchanged. For instance, the Obscurer could agree to never follow a lie with a truth, as in [1].

In section 2 we study whether, given a specific restriction on the Obscurer's inaccuracy, the Seeker can win at all. Then in section 3 we apply the tools we have built to the family of restrictions characterized by forbidden substrings in the series of truths and lies.

## 2 Renyi-Ulam games in general

**Definition 1.** A *lie restriction*  $R$  is a condition on binary strings.

A 0 in a binary string stands for a truth and a 1 stands for a lie. We shall only study lie restrictions  $R$  such that:

- $R$  is *prefix-closed*, that is, every prefix of a string that satisfies  $R$  also satisfies  $R$ , and
- $R$  is *extensible*, that is, every string that satisfies  $R$  can be extended by one letter so as to continue to satisfy  $R$ .

This is a natural requirement. As a string that satisfies a restriction  $R$  is formed over the course of a game,  $R$  needs to be prefix-closed (for, otherwise, the Obscurer has already broken  $R$  on a previous move) and extensible (for, otherwise, the Obscurer would have no legal reply should the Seeker continue to ask questions).

Here are some restrictions that have been looked into. See [2] for details.

- The string  $s$  contains at most  $c$  lies, where  $c$  is a fixed positive integer.
- Every prefix  $p$  of the string  $s$  contains at most  $c \cdot |p|$  lies, where  $c$  is a fixed positive real.
- Every substring  $p$  of the string  $s$  such that  $|p| = a$  contains at most  $b$  lies, where  $a$  and  $b$  are fixed positive integers.
- The string  $s$  does not contain a substring equal to  $p$ , where  $p$  is a fixed string.

**Definition 2.** The *Renyi-Ulam game with lie restriction  $R$*  is played as follows.

There are two players: the Seeker and the Obscurer. The Obscurer thinks of a number between 1 and  $n$ . The Seeker wishes to identify that number.

On each turn, the Seeker asks the Obscurer whether her number belongs to a specific subset of the numbers from 1 to  $n$ . The Obscurer answers either yes or no, and her answer is either true or false. On every turn, the series of truths and lies given by the Obscurer thus far must belong to  $R$ .

Consider the Seeker's state of knowledge by the end of turn  $i$ . For each binary string  $s$  of length  $i$  that satisfies  $R$ , the Seeker knows that if the accuracy of the Obscurer's replies thus far is described by  $s$ , then the Obscurer's number belongs to some subset  $A(s)$  of the numbers from 1 to  $n$  (where  $A(s)$  depends on the particular series of questions asked by the Seeker).

Whenever the Seeker asks whether the Obscurer's number belongs to a set  $Q$ , each  $A(s)$  is split: if the Obscurer answers "yes", then  $A(s)$  is split into  $A(s0) = A(s) \cap Q$  and  $A(s1) = A(s) \cap \overline{Q}$ , and if she answers "no", then  $A(s)$  is split into  $A(s0) = A(s) \cap \overline{Q}$  and  $A(s1) = A(s) \cap Q$ . Furthermore, whenever  $s0$  or  $s1$  does not satisfy  $R$ , the corresponding subset of possibilities drops out of consideration.

The Seeker's goal is to attain a state of knowledge such that all  $A(s)$  but one are empty, the exceptional  $A(s)$  containing exactly one element: the Obscurer's number.

A key observation at this point is that, since the  $A(s)$  are disjoint following each turn, all that matters is the number of elements in each  $A(s)$  rather than which elements precisely comprise it. Therefore, it makes sense to instead consider  $a(s)$ , the number of possibilities for the Obscurer's number that each  $s$  permits.

On turn 0, before the game begins,  $a(\varepsilon) = n$ . On each subsequent turn, the Seeker effectively picks two numbers  $b(s)$  and  $c(s)$  for each  $a(s)$  so that  $a(s) = b(s) + c(s)$ , asks a question whose  $Q$  intersects each  $A(s)$  by  $b(s)$  elements, and arrives at  $a(s0) = b(s)$  and  $a(s1) = c(s)$  if the Obscurer answers "yes" and at  $a(s0) = c(s)$  and  $a(s1) = b(s)$  if she answers "no". Again, whenever  $s0$  or  $s1$  does not satisfy  $R$ , the corresponding number of possibilities drops out of consideration. The Seeker's goal is to make all  $a(s)$  but one equal to zero, with the exceptional  $a(s)$  equal to one (or, equivalently, to bring the sum of all  $a(s)$  down to unity).

This reduction makes it natural to group together descriptors  $s$  that have identical possible futures. That is, if  $s_1$  and  $s_2$  are such that, for all strings  $s'$ , the strings  $s_1s'$  and  $s_2s'$  either both satisfy  $R$ , or both fail to satisfy  $R$ , then there is no need to deal with  $a(s_1)$  and  $a(s_2)$  separately: they can be replaced with their sum,  $a(s_1) + a(s_2)$ , and it can henceforth be treated as any other  $a(s)$ .

This motivates the following series of definitions.

**Definition 3.** A *lie restriction graph*  $G$  for a lie restriction  $R$  is a directed graph  $G$  such that:

- One vertex in  $G$  is called a *start vertex*.
- Every arc in  $G$  is labeled either 0 or 1.
- Every vertex in  $G$  originates at least one and at most two outward arcs. Of those, at most one arc is labeled 0 and at most one arc is labeled 1.
- A binary string  $s$  satisfies  $R$  if and only if there is a path  $l$  in  $G$  such that  $l$  starts from the start vertex and the arc labels along  $l$  trace out  $s$ .

Two strings  $s_1$  and  $s_2$  have identical possible futures iff the paths that they define in  $G$  starting from the start vertex end up on the same vertex. A restriction graph captures the notion of identical possible futures in precisely the same way as this is done in automata theory. More precisely, if  $A$  is a deterministic automaton that recognizes the language of all strings that satisfy  $R$  then removing all non-accept states from  $A$  results in a restriction graph for  $R$ .

Every prefix-closed extensible restriction  $R$  possesses at least one restriction graph  $G$ : the vertices of  $G$  are all strings  $s$  that satisfy  $R$ , and an arc labeled  $a$  in  $G$  always points from a string  $s$  to  $sa$ . Of course, given a prefix-closed extensible restriction  $R$ , there may be many distinct restriction graphs  $G$  for  $R$ .

Consider, then, a fixed restriction  $R$  and a fixed restriction graph  $G$  for  $R$ .

**Definition 4.** A *position* is a mapping from the vertices of  $G$  to the nonnegative integers.

A position describes the Seeker's state of knowledge following a turn. We assume a numbering of the vertices of  $G$  (with 1 being the start vertex) and write  $(a_1, a_2, \dots)$  for the position that maps vertex  $i$  to the integer  $a_i$ .

**Definition 5.** The *support* of a position  $P$  is the number of vertices  $v$  of  $G$  such that  $P(v)$  is nonzero.

All positions that we are about to consider are of finite support.

**Definition 6.** The *weight* of a position  $P$  is the sum of  $P(v)$  over all vertices  $v$  of  $G$ .

The Seeker's goal is to reach a position of unit weight.

**Definition 7.** Let  $P$ ,  $P_1$ , and  $P_2$  be three positions. We say that  $(P_1, P_2)$  is a *split* of  $P$  if there exist two positions  $Q_0$  and  $Q_1$  such that

- For all vertices  $v$  of  $G$ ,  $P(v) = Q_0(v) + Q_1(v)$ .
- For all vertices  $v$  of  $G$ ,  $P_1(v)$  is the sum of  $Q_0(u)$  over all vertices  $u$  of  $G$  such that a 0-arc in  $G$  points from  $u$  to  $v$ .
- For all vertices  $v$  of  $G$ ,  $P_2(v)$  is the sum of  $Q_1(u)$  over all vertices  $u$  of  $G$  such that a 1-arc in  $G$  points from  $u$  to  $v$ .

A split captures the Seeker's capacity to advance to a different position by asking a question. If the Seeker is currently in a position  $P$  and  $(P_1, P_2)$  is a split of  $P$ , then the Seeker can devise a question that, depending on the Obscurer's answer, will advance the game to either  $P_1$  or  $P_2$ .

**Definition 8.** A *strategy-like tree* is a binary tree such that its vertices are positions and every non-leaf vertex  $P$  has two children that form a split of  $P$ . A strategy-like tree for a position  $P$  is a strategy-like tree rooted at  $P$ .

**Definition 9.** A *strategy tree* is a strategy-like tree whose every leaf has weight either zero or one. A strategy tree for a position  $P$  is a strategy tree rooted at  $P$ .

A leaf of zero weight corresponds to a position that cannot occur in a game, and a leaf of unit weight corresponds to a position in which the Seeker has won.

The Seeker wins the Renyi-Ulam game with lie restriction  $R$  from a position  $P$  iff there is a strategy tree for  $P$ . Consequently, the Seeker wins the Renyi-Ulam game with lie restriction  $R$  for  $n$  iff there is a strategy tree for the position  $(n, 0, 0, \dots)$ .

We are done setting up the scaffolding and proceed to give two criteria on winning positions.

**Theorem 1.** *The Seeker wins from all positions iff the Seeker wins from all positions of weight two.*

*Proof.* Suppose that the Seeker wins from all positions of weight two. Then there is a strategy tree for every position of weight two. We will prove that there is a strategy tree for every position by induction on weight.

Let  $P$  be a position such that  $w(P) \geq 3$ . Then either  $P$  has a coordinate greater than or equal to two, or  $P$  has at least two coordinates greater than or equal to one. In both cases,  $P = P_1 + P_2$  for some positions  $P_1$  and  $P_2$  such that  $w(P_1)$  and  $w(P_2)$  are both strictly less than  $w(P)$ .

Let  $T_1$  and  $T_2$  be strategy trees for  $P_1$  and  $P_2$ . Extend  $T_1$  and  $T_2$  to two isomorphic strategy trees  $T'_1$  and  $T'_2$  by adding on sufficiently many leaves of weight zero or one. Furthermore, let  $T'$  be isomorphic to both  $T'_1$  and  $T'_2$  and every vertex of  $T'$  be the sum of the corresponding vertices of  $T'_1$  and  $T'_2$ . Then  $T'$  is a strategy-like tree for  $P$  all of whose leaves have weight at most two. Replace every leaf  $L$  of  $T'$  with a strategy tree for  $L$ , and the result is a strategy tree for  $P$ . □

**Theorem 2.** *The Seeker wins from all positions of the form  $(n, 0, 0, \dots)$  iff the Seeker wins from all positions of the form  $(0, \dots, 0, 2, 0, \dots)$  (unit support, weight two).*

*Proof.* Suppose first that the Seeker wins from all positions of the form  $(n, 0, 0, \dots)$ . Let  $u$  be a vertex in  $G$ . Let  $a$  be the starting vertex: then there is a path  $a-v_1-\dots-v_k-u$  from  $a$  to  $u$ .

Let  $T$  be a strategy tree for the position  $(2^{k+2}, 0, 0, \dots)$ . One of the children  $P_1$  of this position in  $T$  has a  $v_1$ -coordinate of at least  $2^{k+1}$ . One of the children of  $P_1$ ,  $P_2$ , has a  $v_2$ -coordinate of at least  $2^k$ , and so on. Eventually, some vertex of  $T$  has a  $u$ -coordinate of at least 2. This means that the Seeker wins from the position that maps  $u$  to two and all other vertices to zero.

Suppose, then, that the Seeker wins from all positions of the form  $(0, \dots, 0, 2, 0, \dots)$ . We set out to prove that the Seeker wins from all positions of the form  $(0, \dots, 0, n, 0, \dots)$  by induction on  $n$ .

Let  $n \geq 3$  and  $m = \lceil \frac{n}{2} \rceil$ . Let  $P$  be a position of the form  $(0, \dots, 0, n, 0, \dots)$ . Obtain  $P'$  from  $P$  by replacing  $n$  with  $m$  and obtain  $P''$  from  $P'$  by replacing  $m$  with two.

Let  $T''$  be a strategy tree for  $P''$ . Obtain  $T'$  from  $T''$  by multiplying all vertices of  $T''$  by  $m$ . Then  $T'$  is a strategy-like tree for  $P'$  all of whose leaves are either of the form  $(0, \dots, 0, m, 0, \dots)$  or  $(0, 0, \dots)$ .

By induction hypothesis, it is possible to replace every leaf  $L$  of  $T'$  with a strategy tree for  $L$ , thus obtaining a strategy tree for  $P'$ . It follows that the Seeker wins from  $P'$  and, therefore, from  $P$ .  $\square$

The situations discussed in Theorem 1 and Theorem 2 are distinct: there are restrictions  $R$  such that the Seeker wins from all unit-support weight-two positions but not from all weight-two positions. One example of this is given by the restriction graph  $\{(1, 2), (1, 3), (2, 1), (3, 1)\}$  corresponding to the restriction that answers  $2i - 1$  and  $2i$  are either both truths or both lies for all  $i$ .

**Definition 10.** A lie restriction  $R$  is *regular* if the set of all strings that satisfy  $R$  is a regular language.

A restriction  $R$  is regular iff there is a finite restriction graph  $G$  for  $R$ .

**Theorem 3.** Let  $R$  be a regular restriction and  $G$  be a finite restriction graph for  $R$  with  $k$  vertices. Suppose that the Seeker wins for  $n = 2^k$ . Then the Seeker wins for all  $n$ .

*Proof.* As in the first part of the proof of Theorem 2, given that the Seeker wins for all  $n = 2^k$ , we conclude that the Seeker wins from all positions of the form  $(0, \dots, 0, 2, 0, \dots)$ . By Theorem 2, the proof is complete.  $\square$

**Theorem 4.** Let  $R$  be a regular restriction such that the Seeker wins for all  $n$ . Then the Seeker wins in  $O(\log n)$  steps.

*Proof.* Let  $c$  be such that the Seeker wins all positions of the form  $(0, \dots, 0, 2, 0, \dots)$  in at most  $c$  steps and  $h(n)$  be the number of iterations of the transform  $n \rightarrow \lceil \frac{n}{2} \rceil$  that it takes to boil  $n$  down to 1. The construction in the second part of the proof of Theorem 2 shows that there is a strategy tree for the position  $(n, 0, 0, \dots)$  of depth at most  $c \cdot h(n)$ , and this is  $O(\log n)$ .  $\square$

There are non-regular restrictions  $R$  such that the Seeker wins for all  $n$  but cannot do so in a logarithmic number of steps. For instance, construct a restriction graph  $G$  for  $R$  as follows:

- The vertices of  $G$  are  $(a, b)$ , where  $a$  ranges over all positive integers and  $1 \leq b \leq F(a)$  where  $F$  grows very rapidly.
- $(a, 1)$  points to  $(a + 1, 1)$  and  $(a, 2)$  for all  $a$ .
- For all  $b \geq 2$ ,  $(a, b)$  points to itself and, as long as  $b < F(a)$ , to  $(a, b + 1)$ .

We proceed to describe an equivalent form of the criterion given in Theorem 2 which shall prove useful in section 3.

**Definition 11.** The *two-string game* for a lie restriction  $R$  is played between two players as follows. The starting position is given by two binary strings  $u_1$  and  $u_2$ . On each turn, the Obscurer extends one of  $u_1$  and  $u_2$  (that is, appends a character to it) and the Seeker replies by extending the other. The Seeker's goal is that one of the two strings ceases to satisfy  $R$  (it does not matter whether this happens following a move by the Seeker or the Obscurer) and the Obscurer's goal is to prevent this.

We shall freely identify strings that satisfy a restriction  $R$  with vertices in a restriction graph  $G$  for  $R$ . Moves in the two-string game can thus also be regarded as transitions within  $G$ .

**Theorem 5.** *The Seeker wins the Renyi-Ulam game with lie restriction  $R$  from a position  $P$  of weight two iff the Seeker wins the two-string game for  $R$  from a starting position given by the  $P$ -nonzero vertices of the lie restriction graph  $G$  (or, should there be just one such vertex  $v$ , by two copies of  $v$ ).*

*Proof.* Let  $P_{x,y}$  be the position of weight two with nonzero coordinates  $x$  and  $y$  (if  $x \neq y$  then both the  $x$  and  $y$  coordinate of  $P_{x,y}$  equal one, and if  $x = y$  then the  $x$ -coordinate of  $P_{x,y}$  equals two).

Obtain  $H$  from  $G$  by adjoining an additional *error vertex*  $e$  such that:

- for every vertex  $x$  of  $G$  and every label  $a$  such that in  $G$  no arc labeled  $a$  originates at  $v$ , in  $H$  an arc labeled  $a$  points from  $v$  to  $e$ , and
- in  $H$  two arcs labeled 0 and 1 point from  $e$  to  $e$ .

Then in  $H$  every vertex has out-degree two.

Consider  $P = P_{u,v}$ . Let, in  $H$ ,  $a$  and  $b$  be the children of  $u$  and  $c$  and  $d$  be the children of  $v$ .

Suppose that the Obscurer wins the Renyi-Ulam game with restriction  $R$  from  $P_{u,v}$ . There are only two possibilities for the children of  $P_{u,v}$  in a strategy-like tree:  $P_{a,c}$  and  $P_{b,d}$ , or  $P_{a,d}$  and  $P_{b,c}$ . Since  $P_{u,v}$  is a (Renyi-Ulam) win for the Obscurer, (at least) one position in the first pair and (at least) one position in the second pair are also (Renyi-Ulam) wins for the Obscurer. Without loss of generality, those are  $P_{a,c}$  and  $P_{a,d}$ : in this case, we say that  $a$  is a key vertex for  $P_{u,v}$ .

Consider the two-string game for  $R$  played from  $P_{u,v}$ . By always moving to the key vertex for the current position, the Obscurer attains that the position following the Seeker's move is always a (Renyi-Ulam) win for the Obscurer. This ensures that none of the two strings ever reaches the error vertex.

For the converse, let  $P_{u,v}$  be an arbitrary position such that the Obscurer wins the two-string game for  $R$  starting from  $P_{u,v}$ . Suppose that transitioning

from  $u$  to  $a$  is a (two-string) winning move for the Obscurer. Then, in the Ulam-Renyi game with lie pattern  $R$ , it is possible that the Obscurer replies to the Seeker's question in such a way that the resulting position has a nonzero  $a$  coordinate. Since, in the two-string game, no string ever reaches the error vertex, this ensures that no position of weight one ever occurs.  $\square$

**Theorem 6.** *The Seeker wins the Renyi-Ulam game with lie restriction  $R$  from all positions of the form  $(n, 0, 0, \dots)$  iff the Seeker wins the two-string game for  $R$  from all starting positions such that the two initial strings  $u_1$  and  $u_2$  are equal.*

*Proof.* This is a strightforward corollary of Theorem 2 and Theorem 5.  $\square$

Notice that, when  $G$  is finite, the number of essentially different positions in the two-string game for  $R$  is bounded by the square of the number  $k$  of vertices in  $G$ . Consequently, when  $k$  is small, it is feasible to determine who has a winning strategy in the two-string game for  $R$  by means of direct computation. We shall make frequent use of this technique in the following section.

### 3 Forbidden substrings

Let  $S$  be a set of binary strings and  $R = r(S)$  be the restriction that a binary string  $s$  does not contains a substring in  $S$ . Given  $S$ , who wins in a Renyi-Ulam game with restriction  $R$ ?

The case  $|S| = 1$  was first studied, and solved completely, by Czyzowicz, Lakshmanan and Pelc in [1]. Theorem 7 below is the main result in their paper.

**Definition 12.** We say that  $S_1$  is *expanded* by  $S_2$  if  $S_1$  contains a substring of every string in  $S_2$ .

In this case,  $R_1 = r(S_1) \subseteq R_2 = r(S_2)$  and  $R_1$  is more restrictive on the Obscurer's capacity to mislead the Seeker than  $R_2$ . This means that, if the Obscurer wins for  $R_1$ , then she also wins for  $R_2$ . Conversely, if the Seeker wins for  $R_2$ , then he also wins for  $R_1$ .

We set out to find a natural sufficient condition for the Obscurer to win. By Theorem 6, it suffices to find a sufficient condition for the Obscurer to win the two-string game for  $R$  from some starting position comprised of two identical strings.

**Definition 13.** The *risk* of a string  $p$  is the length of the longest suffix of  $p$  that is a prefix of an element of  $S$ . When  $p$  is of risk  $l$  and the  $l$ -suffix of  $p$  is an  $l$ -prefix of  $q \in S$ , we say that  $p$  is of risk  $l$  *via*  $q$ . The *risk* of a position  $P$  in the two-string game for  $R$  is the maximum of the risks of the two strings that comprise  $P$ .



Notice that extending a string increases its risk by at most one.

Suppose that all strings in  $S$  are of length at least  $l + 1$  and that (starting from a position of risk at most  $l - 1$ , say, two empty strings) the Obscurer can play so that, following her every move, the risk of the position is at most  $l - 1$ . Then the Obscurer wins.

What is a natural sufficient condition for that?

**Definition 14.** The  $l$ -flip of a string  $p = a_1a_2 \dots a_m$  of length  $m \geq l + 1$  is the string  $a_1a_2 \dots a_l\overline{a_{l+1}}$ .

Suppose that, for some  $l \geq 1$ , all  $(l - 1)$ -flips and all  $l$ -flips of elements of  $S$  are of risk at most  $l - 1$ . Then the Obscurer can keep risk following her moves below or at the  $l - 1$  level by always spoiling the riskier string. This is the condition we were looking for, and its negation is a necessary condition for the Seeker to win.

**Lemma 1.** *For the Obscurer to win the Renyi-Ulam game with lie restriction  $R = r(S)$ , it is sufficient that there is a positive integer  $l$  such that all  $(l - 1)$ -flips and all  $l$ -flips of elements of  $S$  are of risk at most  $l - 1$ .*

*For the Seeker to win the Renyi-Ulam game with lie restriction  $R = r(S)$ , it is necessary that, for all  $l \geq 1$ , there is either some  $(l - 1)$ -flip or some  $l$ -flip of an element of  $S$  of risk at least  $l$ .*

Let us test the strength of this criterion on a known result.

**Theorem 7.** *When  $S = \{s_1\}$ , the Seeker wins iff  $s_1$  is one of 0, 1, 01, and 10.*

*Proof.* Notice first that the  $l$ -flip of  $s_1$  is not a prefix of  $s_1$  so that its risk does not exceed  $l$ .

Suppose that  $|s_1| \geq 3$ ,  $s_1 = a_1a_2 \dots a_m$ , and, in Lemma 1, set  $l = 2$ . Then the 2-flip of  $s_1$  must be of risk 2, which gives us  $a_2\overline{a_3} = a_1a_2$ .

Set, then,  $l = 1$  in Lemma 1: analogously, the 1-flip of  $s_1$  needs to be of risk 1, giving us  $\overline{a_2} = a_1$ , a contradiction.

It follows that  $|s_1| \leq 2$  and direct computation yields the claim.  $\square$

We turn to a complete description of all  $S = \{s_1, s_2\}$  such that the Seeker wins. In the series of lemmas that follow, we assume that  $S = \{s_1, s_2\}$  and  $|s_1| \leq |s_2|$  except where noted otherwise.

**Lemma 2.** *If the Seeker wins, then the strings  $s_1$  and  $s_2$  do not share a common substring  $p$  equal to 00 or 11 or of length at least three.*

*Proof.* For, otherwise  $\{p\}$  is expanded by  $S$  and Theorem 7 gives us a contradiction.  $\square$

**Lemma 3.** *If the Seeker wins, then the length of  $s_1$  does not exceed four.*

*Proof.* Suppose that  $5 \leq |s_1| \leq |s_2|$  and apply Lemma 1 with  $l = 4$ .

Suppose that a 3-flip of  $s_1$  is of risk 4. It cannot be so via  $s_1$  because of the definition of a flip, nor can it be so via  $s_2$  because of Lemma 2: a contradiction. Analogously, a 3-flip of  $s_2$  cannot be of risk 4.

Suppose, then, that a 4-flip of  $s_1$  is of risk at least 4. It cannot be so via  $s_2$  because of Lemma 2. Therefore, it must be so via  $s_1$ , which gives us  $s_1 = 00001 \dots$  modulo inverting all characters. (The case when a 4-flip of  $s_2$  is of risk 4 is treated analogously.)

Next we apply Lemma 1 with  $l = 3$ .

As above, no 2-flip of either string may be of risk 3 and no 3-flip of either string may be of risk 3 via the other (it is important at this point that we know that  $s_1$  starts on a series of equal characters so that Lemma 2 applies).

If a 3-flip of  $s_1$  is of risk 3 via  $s_1$ , we arrive at a contradiction. Therefore, a 3-flip of  $s_2$  is of risk 3 via  $s_2$  and  $s_2 = 0001 \dots$  or  $s_2 = 1110 \dots$ .

When  $s_2 = 0001 \dots$ , Lemma 2 gives us a contradiction. When  $s_2 = 1110 \dots$ , Lemma 1 gives us a contradiction with  $l = 2$ . □

Notice that we could have obtained the same result by testing all pairs of length-5 strings directly. The above argument, however, generalizes: see Theorem 10.

**Lemma 4.** *When  $|s_1| = 4$ , the Seeker wins for the following sets  $S$ :  $\{0001, 1011\}$ ,  $\{0001, 1101\}$ ,  $\{0010, 0111\}$ ,  $\{0011, 0101\}$ ,  $\{0100, 0111\}$ , and the sets obtained from them by inverting all characters.*

*Proof.* Direct computation shows that there are no other  $S$  such that the Seeker wins when  $|s_1| = 4$  and  $|s_2| \leq 5$ . Since every  $S$  such that  $4 = |s_1| \leq |s_2|$  expands some  $S$  such that  $4 = |s_1| \leq |s_2| \leq 5$ , the claim follows. □

**Lemma 5.** *When  $s_1 = 000$ , the Seeker wins for the following  $s_2$ : 011, 101, 110, 1011, and 1101.*

*Proof.* Analogous to the proof of Lemma 4. □

Next comes a brief detour whose purpose will become evident later.

**Lemma 6.** *When  $d \geq 2$  and  $S = \{01^{kd}0 \mid k \geq 0\}$ , the Obscurer wins from the empty-string starting position.*

*Proof.* If  $u_1$  does not contain any 0s, set  $h_1 = \infty$ . Otherwise, set  $h_1$  equal to the remainder of the number of 1s since the last 0 in  $u_1$  upon division by  $d$ . Define  $h_2$  analogously for  $u_2$ .

We say that a position is *safe* if it is the Seeker to move, both  $h_1$  and  $h_2$  differ from 0, and, additionally, if  $h_i + 1 = h_{\sim i} \neq \infty$ , then the Seeker is to extend  $h_{\sim i}$ . (Here,  $\sim 1 = 2$  and  $\sim 2 = 1$ .)

The Obscurer's strategy is to always write down a 1 so as to make the position safe. We proceed to show that this is indeed possible by means of induction and casework.

The Obscurer's opening move results in a safe position with  $h_1 = h_2 = \infty$ .

Suppose that the Seeker is to move in a safe position with  $h_1 \leq h_2$ . We write Player:  $i \circ a$  for Player extending  $u_i$  by  $a$ . Here follows a list of all possible situations together with viable replies for the Obscurer. Immediately forced timelines of the form

Seeker:  $i \circ$  any leading to  $h_i = 0$ , Obscurer:  $i \circ 1$   
are omitted for brevity.

- $h_1 + 1 < h_2 < d - 1$ , Seeker:  $i \cdot 1$ , Obscurer:  $\sim i \cdot 1$
- $h_1 + 1 = h_2 < d - 1$ , Seeker:  $2 \cdot 1$ , Obscurer:  $1 \cdot 1$
- $h_1 = h_2 < d - 1$ , Seeker:  $i \cdot 1$ , Obscurer:  $\sim i \cdot 1$
- $h_1 + 1 < h_2 = d - 1$ , Seeker:  $1 \cdot 1$ , Obscurer:  $1 \cdot 1$
- $h_1 + 1 = h_2 = d - 1$
- $h_1 = h_2 = d - 1$
- $h_1 < h_2 = \infty$ , Seeker:  $i \cdot 1$  not leading to  $h_1 = 0$ , Obscurer:  $2 \cdot 1$
- $h_1 = h_2 = \infty$ , Seeker:  $i \cdot 1$ , Obscurer:  $i \cdot 1$ .

All of those lead back to a safe position, and we are done. □

A straightforward corollary is

**Lemma 7.** *When  $k \geq 2$  and  $S = \{00, 01^k 0\}$ , the Obscurer wins from the empty-string starting position.*

Let  $R_\#$  be the string set  $r(\{01^k 0 \mid k = 0 \text{ or } k \geq 2\})$ . Then

$$R_\# = \{1^a \underbrace{010 \dots 10}_b 1^c \mid a, b, c \geq 0, b = 0 \text{ or } b \text{ is odd}\}.$$

**Lemma 8.** *When  $S = \{00, s_2\}$  with  $s_2$  in  $R_\#$ , the Seeker wins.*

*Proof.* Notice first that the Obscurer must only write down 1s. For, to Obscurer:  $i \circ 0$ , the Seeker replies  $\sim i \circ 0$  and wins on the following turn.

Also, Seeker:  $i \circ 0$  forces Obscurer:  $i \circ 1$  – for, if the Obscurer replies in any other way, the Seeker wins on the turn.

Take, then,  $s_2 = 1^a \underbrace{010 \dots 10}_b 1^c$ . The Seeker's strategy is as follows.

- If  $b = 0$ , the Seeker writes down 1s only until  $s_2$  occurs.
- From this point on,  $b \geq 1$ . The Seeker writes down 1s only until he is to extend some  $u_i$ , say  $u_1$ , that ends in at least  $a$  1s.
- The Seeker loops over [Seeker:  $1 \circ 0$ , Obscurer:  $1 \circ 1$ , Seeker:  $2 \circ 0$ , Obscurer:  $2 \circ 1$ ]  $\frac{b+1}{2}$  times.
- The Seeker extends  $u_1$  by 1s only and  $u_2$  by 0s only until  $s_2$  occurs in  $u_1$ .

□

Lemma 7 and Lemma 8 give us

**Lemma 9.** *When  $S = \{00, s_2\}$ , the Seeker wins iff  $s_2 \in R_\#$ .*

**Lemma 10.** *When  $s_1 = 001$ , the Seeker wins iff  $s_2 \in R_\#$ .*

*Proof.* By Lemma 9, this condition is necessary. We proceed to show that it is also sufficient.

By Lemma 9, the Seeker can play so that either  $00$  or  $s_2$  occurs following a move by the Seeker. If the latter, we are done. If the former, then without loss of generality  $00$  has occurred in  $u_1$ . The following scenario tree,

- Obscurer:  $1 \circ 0$ , Seeker:  $2 \circ 0$ ,
  - Obscurer:  $1 \circ 0$ , Seeker:  $2 \circ 0$ , Obscurer: any, Seeker: any  $\circ 1$
  - Obscurer:  $2 \circ$  any, Seeker:  $1 \circ 1$
- Obscurer:  $2 \circ$  any, Seeker:  $1 \circ 1$

(timelines such that the Seeker wins on the Obscurer's move are omitted for brevity) shows that the Seeker wins in at most two more turns.

□

**Lemma 11.** *When  $s_1 = 010$ , the Seeker wins for the following  $s_2$ :  $001$ ,  $011$ ,  $100$ ,  $110$ ,  $111$ ,  $0011$ ,  $0111$ ,  $1100$ ,  $1110$ ,  $00111$ , and  $11100$ .*

*Proof.* Analogous to the proofs of Lemma 4 and Lemma 5.

□

**Lemma 12.** *When  $s_1 = 100$ , the Seeker wins iff  $s_2 \in R_\#$ .*

*Proof.* By Lemma 9, this condition is necessary. We proceed to show that it is also sufficient.

Initially, the Seeker ensures that both  $u_1$  and  $u_2$  contain a 1. Here follows one way to achieve this: without loss of generality, the Seeker is to extend  $u_1$ . He plays  $1 \circ 1$ , then proceeds to play  $1 \circ 0$  and  $1 \circ 0$  until he either wins by completing

100 in  $u_1$  or the Obscurer interrupts him, giving him the opportunity to play  $2 \circ 1$ .

Following this, by Lemma 9, the Seeker plays so that either 00 or  $s_2$  occurs. If the latter, we are done. If the former, then we are done as well because a 1 preceding 00 yields an occurrence of 100.

□

All remaining  $|s_1| \leq 3$  cases are reduced to the ones described above by inverting all characters in  $S$ .

All cases such that  $|s_1| \leq 2$  are handled by Theorem 7, Lemma 9, and, for  $s_2 = 11$ , inverting all characters.

This completes the classification, and the following theorem summarizes it.

**Theorem 8.** *The complete list of all two-element sets of binary strings  $S$  such that the Seeker wins the Ulam-Renyi game with lie restriction  $R = r(S)$  is given by all sets expanded by a single-element Seeker-win set as listed in Theorem 7, the 28 sporadic pairs*

$s_1 = 000, s_2 = 101, 1011, 1101,$   
 $s_1 = 010, s_2 = 111, 0011, 0111, 1100, 1110, 00111, 11100,$   
 $s_1 = 101, s_2 = 0001, 0011, 1000, 1100, 00011, 11000,$   
 $s_1 = 111, s_2 = 0010, 0100,$   
 $\{0001, 1011\}, \{0001, 1101\}, \{0010, 0111\}, \{0010, 1110\}, \{0011, 0101\},$   
 $\{0100, 0111\}, \{0100, 1110\}, \{1000, 1011\}, \{1000, 1101\}, \{1010, 1100\},$

and the six infinite families

$s_1 \in \{00, 001, 100\}, s_2 \in R_{\#}$  and  $s_1 \in \{11, 011, 110\}, s_2 \in R_{\#}^{inv}$ .

A dual classification is given by the minimal (with respect to shrinking any element to a proper substring) Obscurer-win two-element  $S$ . It consists of 16 sporadic pairs,

$s_1 = 000, s_2 = 010, 111, 11011,$   
 $s_1 = 010, s_2 = 101, 0110, 1001, 1111, 01110,$   
 $s_1 = 101, s_2 = 111, 0000, 0110, 1001, 10001,$   
 $s_1 = 111, s_2 = 00100,$   
 $\{0011, 1010\},$  and  $\{0101, 1100\},$

together with two infinite families,

$\{00, 01^k 0\}$  and  $\{11, 10^k 1\}, k \geq 2.$

By Theorem 8, the language formed by all strings  $s_1 * s_2$  (where  $*$  is a new character distinct from 0 and 1) such that the Seeker wins for  $S = \{s_1, s_2\}$  is regular. As the following theorem shows, an analogous claim does not hold for three-element  $S$ .

**Theorem 9.** *When  $S = \{00, 01^m 0, 01^n 0\}$ ,  $m, n \geq 1$ , the Seeker wins the Renyi-Ulam game with lie restriction  $R = r(S)$  iff  $m$  and  $n$  are coprime.*

*Proof.* Suppose first that  $(m, n) = d > 1$ . By Lemma 6, the Obscurer wins from the empty-string starting position.

Suppose, then, that  $(m, n) = 1$ . We proceed to show that the Seeker wins by induction on  $\max\{m, n\}$ .

When one of  $m$  and  $n$  equals 1, the claim follows by Theorem 8. Assume, from this point on, that  $2 \leq m < n$ .

As in the proof of Lemma 8, the Obscurer must only write down 1s, and Seeker:  $i \circ 0$  forces Obscurer:  $i \circ 1$ .

The Seeker begins by writing down 1s only for the first  $n + 1$  turns. This ensures that, from turn  $n + 2$  on, the game proceeds as if from the empty-string starting position.

By induction hypothesis, the Seeker wins for  $S' = \{00, 01^m 0, 01^{n-m} 0\}$ . Let the Seeker follow a strategy that wins for  $S'$  from the empty-string starting position. Then there must come a moment when the Seeker is to move, the Seeker has not won yet (for  $S'$ ) and cannot win on this turn (for  $S'$ ), and the Seeker can play so as to win on the next turn (for  $S'$ ) regardless of the Obscurer's reply.

At this point, one turn short of the final stroke, the Seeker abandons the  $S'$  strategy and focuses back on his original goal, winning for  $S$ .

In a position  $u_1, u_2$  such that it is the Obscurer to move, the Seeker has not won yet (for  $S'$ ), and the Seeker wins on the turn (for  $S'$ ), both of  $u_1$  and  $u_2$  must necessarily end in one of  $0, 01^m$ , or  $01^{n-m}$ .

This means that, without loss of generality, at present  $u_1$  ends in one of  $0, 01^m$ , or  $01^{n-m}$  and the Seeker is to extend  $u_2$ .

The Seeker plays  $2 \circ 0$ , forcing Obscurer:  $2 \circ 1$ . If  $u_1$  ends in  $0$  or  $01^m$ , the Seeker wins on his next move.

Otherwise,  $u_1$  ends in  $01^{n-m}$ . The Seeker then writes down 1s only on his next move as well as on the  $m - 1$  turns that follow it. This brings the Seeker and the Obscurer to a point when it is the Obscurer to move,  $u_1$  ends in  $01^n$ , and  $u_2$  ends in  $01^m$ . The Seeker wins on the turn.

□

Theorem 9 hints that there may be a lot more to uncover on forbidden string sets  $S$  with  $|S| \geq 3$ . We conclude with a generalization of Lemma 3 that applies to all  $|S| \geq 2$ .

**Theorem 10.** *When  $|S| = k \geq 2$  and the Seeker wins the Renyi-Ulam game with lie restriction  $R = r(S)$ , the length of the shortest string in  $S$  does not exceed  $4k - 4$ .*

*Proof.* We proceed by induction on  $k$ . When  $k = 2$ , the claim holds by Lemma 3.

Let  $f(k) = 4k - 4$ , and suppose that all elements of  $S$  are of length at least  $f(k - 1) + 5$ .

No two elements of  $S$  share a common substring of length at least  $f(k-1)+1$ . For, otherwise, we could obtain  $S'$  from  $S$  by replacing them with this common substring and, since  $S'$  is expanded by  $S$ , arrive at a contradiction by the induction hypothesis.

Apply, then, Lemma 1 with  $l \geq f(k-1) + 2$ .

Suppose first that an  $(l-1)$ -flip of  $s_i$  is of risk  $l$ . Then it cannot be so via  $s_i$  because of the definition of a flip, nor can it be so via  $s_j$ ,  $j \neq i$ , because that would give us a long common substring: a contradiction.

Suppose, then, that an  $l$ -flip of  $s_i$  is of risk at least  $l$ . It cannot be so via  $s_j$ ,  $j \neq i$ , because that would give us a long common substring. Therefore, it must be so via  $s_i$ , which gives us  $s_i = 0^l 1 \dots$  or  $s_i = 1^l 0 \dots$ .

Apply the previous argument with  $l$  set successively to  $f(k-1) + 2$ ,  $f(k-1) + 3$ , and  $f(k-1) + 4$  to obtain three distinct elements of  $S$  starting on exactly  $f(k-1) + 2$ ,  $f(k-1) + 3$ , and  $f(k-1) + 4$  identical characters. Two of those elements share a common prefix of length at least  $f(k-1) + 2$ , and this gives us a long common substring: a contradiction.

Therefore, at least one string in  $S$  is of length at most  $f(k-1) + 4 = f(k)$ .  $\square$

## References

- [1] Jurek Czyzowicz, K.B. Lakshmanan, and Andrzej Pelc. Searching with a forbidden lie pattern in responses. *Information Processing Letters*, 37(3):127 – 132, 1991.
- [2] Andrzej Pelc. Searching games with errors—fifty years of coping with liars. *Theoretical Computer Science*, 270(1):71 – 109, 2002.
- [3] Alfred Renyi. On a problem in information theory. *Magyar Tud. Akad. Mat. Kutato Int. Kozl*, 6:505–516, 1961.
- [4] Stanislaw Ulam. *Adventures of a Mathematician*. Scribner, 1976.